HUMAN GAIT RECOGNITION BASED ON GAIT ENERGY IMAGE

Macwan Kiran

Assistant Professor
Computer Science & Engineering,
Parul University, Vadodara, India.

Abstract: Biometric system can be used to recognize a person. The biometric system is needed in many key areas such as banking, airport, criminal cases, security purpose etc. Many more colleges and in our Parul university also uses this same system for their employees. Biometric system can recognize person's fingerprints, handprints and it also recognize person's eye (retina) and face. Gait addresses the problem of walking behavior. Gait identification plays a key role in video surveillance and security applications. Gait reorganization can address lack of walking style of person. Gait biometric technology identify a human by the manner, they walk. Based on body movements gait recognition has give a new idea and motivation over the past few years due to the launch of low cost depth cameras. Gait identifies human walking features for individual recognition. In gait recognition, normally, gait feature such as Gait Energy Image (GEI) is extracted from one full gait cycle. Due to some circumstances, such a full gait cycle might not be available due to occlusion. So a new method to identify individuals from gait feature when a few frame(s) is available. I propose a deep learning approach to transform incomplete GEI to the corresponding complete GEI. These transformation is done by several methods and then combining these as a uniform model.

***IndexTerms* – GAIT, GEI.**

I. INTRODUCTION

Human biometric identification systems are used in many organizations such as banking, airport, criminal cases, security purpose etc. Biometric is categorized into two parts namely physiological and behavioral characteristics. Biometric can be used to identify and authenticate individual person based on physiological and behavioral characteristics. In Physiological characteristics area contains the structure of the body. Few examples like fingerprint, Iris scan, footprint etc. Behavioral characteristics are contains to behavior patterns of a person. Few examples like gait, speech patterns, signature etc. The gait cycle is defined as the time interval between the exact same repetitive steps of walking that generally starts when one foot is in contact with the ground. Since last two decades, many gait-based human identification techniques have been proposed. The main observation in all these methods is that a full gait cycle of individuals is available, which is a strong assumption in video surveillance applications where occlusion occurs a lot and a person might be observed in only a few frames. From full Gait cycle, a simple and effective gait representation, namely Gait Energy Image (GEI) is computed by averaging some images of a walking person. This standard gait feature has been broadly used alone or in combination with other features in gait recognition systems. Here a gait-based method to identify a person from a few frames, it is incomplete gait cycle here first generate an incomplete GEI. Next, it train a Fully Convolution Neural Network (FCN) which gets the computed incomplete GEI (average of a few frames) as input and outputs the reconstructed complete GEI. This alteration is done in a progressive way, which means first several auto-encoders are trained as small range repressors. Next the concealed layers of these auto encoders are combined to have a uniform end-to-end network. Here, we have used GEI feature, since it is widely used in many applications. However, the proposed approach can be applied on other gait features such as gait entropy image or gait flow image. The conducted experiments confirm that the proposed network can successfully reconstruct a complete GEI. This can improve gait recognition performance when the input data is not completely available. Therefore, it has great possible in real scenes.

Related Work:

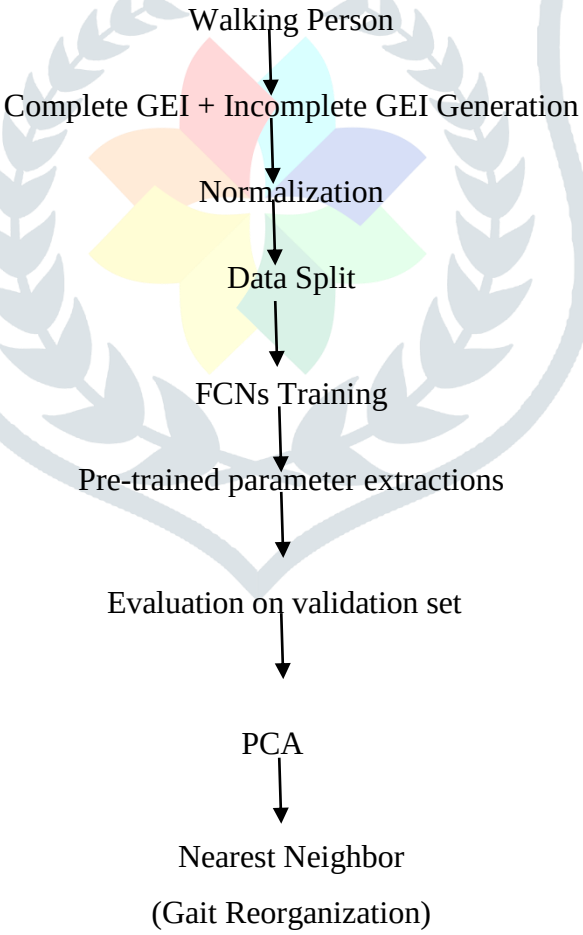
In recent years, various techniques have been proposed for human identification by gait. These techniques can be divided as model-based and model-free approaches. Sundaresan proposed a hidden Markov models (HMMs) based framework for individual recognition by gait. . Huang extend the template matching method to gait recognition by combining transformation based on canonical analysis and eigen space transformation for feature selection. Kshitiz Varma, Sanjeev Sharma(2015) used the front view silhouette data for gait recognition using PCA algorithms. Worapan Kusakunniran, Qiang Wu, Jian Zhang(2013) proposes a new framework to construct a new view-invariant feature for cross-view gait recognition. Procrustes Shape Analysis (PSA) is proposed and applied on a sequence of the normalized gait silhouettes. It is used to extract a new view- invariant gait feature. . The GEI is generated by simply averaging the gait silhouettes over one gait cycle. Other appearance-based feature representations include chrono-gait image (CGI), gait flow image (GFI), and gait entropy image (GEnI). The CGI is a temporal template in which the temporal information among gait frames is encoded by a color mapping function. The CGI is obtained by composing the color encoded gait contour images in a gait cycle. The GFI is generated using optical flow to aggregate gray scale contours over one gait cycle. The GEnI represents the randomness of pixel values in the silhouettes image sequence. In addition, there are also some gait feature representations based on the GEI including enhanced gait energy image (EGEI) , and masked gait energy image (MGEI) . The EGEI retains the dynamic region information to enhance the GEI feature. The MGEI extracts the dynamic information of a gait sequence based on computed entropy. The MGEI is computed based on the GEI and the entropy of every pixel in order to create a binary feature selection mask of GEI. Iwama et. al show that the GEI is more effective in gait recognition than other gait feature representations. The performance of gait recognition is affected by factors such as clothing, carrying objects, view variances, and more importantly, occlusion. Many invariant gait feature representations have been developed, especially for generating view invariant gait features. Deep learning-based approaches execute very well in almost every image and video giving out application, including gait recognition. Recently, extracting useful and view-invariant gait features using deep learning has gained major attention. For example, an invariant feature extraction method has been presented

in , where the authors propose to train a convolution neural network (CNN) using different gait data, including GEIs under different views and different carrying conditions, to extract invariant gait features. These features are learned in a progressive way using multi-stacked auto-encoders. Shiraga et al. propose a CNN called GEINet to extract view-invariant gait features. As a memory-based recognition system, Liu et al. extract 2D positions of joints using the migratory articulated human detection method. Then, a recurrent neural network (RNN) is used for gait recognition. Chao Yan et al. propose to use a CNN and a multi-task learning model (MLT) to identify human gait and to predict multiple human attributes (like view and walking condition) simultaneously. Munif Alotaibi et. al present a gait recognition method using a specialized CNN. They developed a specialized deep CNN architecture that consists of multiple convolutional and sub-sampled layers. This structure is insensitive to earlier mentioned covariant factors which lead to degradation of recognition rates. All mentioned methods compute gait features from one full gait cycle. According to my knowledge, no approach has addressed the problem of gait recognition from an incomplete gait cycle yet. A fully convolutional network (FCN) is a type of deep neural network which can perform image-to-image transformation. It has been successfully applied in image and video super resolution semantic segmentation, and object detection. Dong et al. propose a FCN for single image super resolution. This approach can learn an end-to-end mapping between a low resolution image and a high resolution image, and outperforms traditional sparse coding-based single image super-resolution methods. Our proposed model for incomplete GEI to complete GEI transformation is based on FCN architecture.

Approach

Proposed approach for recognition of individuals from unfinished gait cycle is skilled in two steps. First, an incomplete GEI is computed from available silhouette images by simply averaging them. Second, reconstruct a complete GEI (RC-GEI) from the incomplete GEI (IC-GEI) using the proposed model. An incomplete GEI for a specific person varies based on the different frames and the starting point from which it is generated. This means, in an extreme case, there might be only one or two frames of a gait cycle available. In such case, Input GEI and targeted GEI used for handling the large difference between input (incomplete GEI) and target (complete GEI) is not easy for a single auto encoder. So, train an end-to-end auto-encoder called ITCNet (incomplete to complete GEI network) to do this transformation gradually. More specifically, this end-to-end network is composed of the hidden layer of 9 FCNs, each responsible for mapping in a small range; 1/10 of the gait cycle interval. For instance, if the gait cycle is 30 frames long, the transformation range would be 3 time steps; i.e. the first FCN transforms 1f-GEI to 3f-GEI; the second one transforms 3f-GEI to 6f-GEI, and so on (if-GEI denotes a GEI generated from i frames).

Workflow of the proposed approach



Gait Energy Image

Here only consider individual recognition by activity- specific human motion, i.e., regular human walking, which is used in most current approaches of individual recognition by gait. Regular human walking can be considered as complete a one cycle where human motion repeats at a stable frequency. While some gait recognition approaches extract features from the correlation of all the frames in a walking sequence without considering their order, other approaches extract features from each frame and compose a feature sequence for the human walking sequence. During the recognition procedure, these approaches either match the statistics collected from the feature sequence, or match the features between the corresponding pairs of frames in two sequences that are time-normalized with respect to their cycle lengths. The fundamental assumptions made here are: 1) the order of poses in human walking cycles is the same, i.e., limbs move forward and backward in a similar way among normal people, and

2) differences exist in the phase of poses in a walking cycle, the extend of limbs, and the shape of the torso, etc. Under these assumptions, it is possible to represent the spatiotemporal information in a single 2D gait template instead of an ordered image sequence. The gait energy image (GEI) is obtained by averaging the binary silhouette of human body over one gait cycle Let f (x, y) denotes the binary value of a pixel in position (x, y) at time t, and N is the number of frames in one gait cycle.

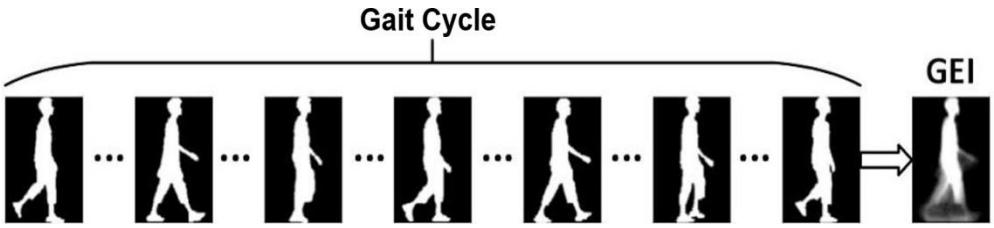


Figure: GEI as a gait feature is computed by averaging gait silhouettes over one gait cycle.

Structure of FCNs:

Propose an incremental GEI reconstruction approach using FCNs that each single FCN enhances the quality of input GEI. Since the gait cycle length depends on the frame rate and is different from a dataset to another dataset, we consider the partial transformations every 1/10 of the gait cycle length. The first FCN transforms a GEI generated from the first 1/10 of the gait cycle to the GEI corresponds to the first 2/10 of the gait cycle. Similarly, the other FCNs enhance their incomplete GEI by predicting the information of the following 1/10 gait cycle. The structures of the all FCNs are the same, but they are trained on different types of GEI. The architecture of one FCN is shown in Fig. 3. Each FCN consists of two parts; the encoder (convolutional) part and the decoder (deconvolutional) part. The encoder part contains three convolutional layers. Each convolutional layer is followed by a Rely function, which serves as a nonlinear activation function, and a pooling layer that down samples the feature maps. After these layers, batch normalization and dropout techniques are used to accelerate convergence and avoid over fitting.

ITCNet: Incomplete to Complete GEI Convertor

In this method, the complete GEI restoration is done in a progressive way (i.e. various types of incomplete GEIs are gradually converted to the complete GEI). The several trained auto-encoders are stacked together to have one end-to-end network (ITCNet). The input of the ITCNet could be any type of incomplete GEI, and the target is the corresponding complete GEI.

Training: As training data, generate multiple incomplete GEIs from different numbers of frames in a gait cycle. Regarding the number of frames, we might have 4 or 8, or in an extreme case there might only be 1 available. It can be generated different incomplete GEIs from different frame numbers N and with different starting frames M for training the Systematic risk is the only independent variable for the CAPM and inflation, interest rate, oil prices and exchange rate are the independent variables for APT model.

Experiments: Here the proposed method was evaluated on two commonly used gait datasets namely OULP & Class B.

OULP:

The OU-ISIR Large Population (OULP) dataset is one of the most recent and popular datasets for gait recognition due to its significant amount of recorded subjects. This dataset provides gait silhouettes of 4007 subjects in size 128 × 88 pixels, including 2135 males and 1872 females with ages ranging from 1 to 94 years old. Each subject in the OULP dataset is recorded under four different observation angles: 55◦, 65◦, 75◦, 85◦. The OULP dataset contains two sections A and B. Section A includes a set of two sequences (gallery and probe sequences) of subjects and is suitable for evaluating the gait recognition performance under constant normal walking conditions, and section B includes a set of subjects with only one sequence and is used for gender-based gait classification.

Class B:

The Institute of Automation from the Chinese Academy of Sciences (CASIA) pro- vided three datasets for gait recognition. The Casia-B dataset is one of the most widely used datasets in recent gait recognition approaches, especially for view-invariant recognition methods, because the data in Casia-B has 11 different viewing angles (0◦, 18◦, ..., 180◦) and different walking scenarios (normal walking, wearing coats or carrying bags). The large number of subjects and the various views of gait data in dataset B make the gait recognition tasks challenging. The normal walking gait sequence includes 124 subjects, with one gallery set and five probe sets for each subject.

Need of Work

In our Parul University and many more organization can recognize their employees by the using of fingerprint, face reorganization. If we can identify person from his/her walking style so it is better solution to authenticate any person. During covid 19 social distancing is necessary so this solution is also helpful in different organization. For medical purpose we can also save human beings walking posture and we can identify from his/her walking style that he/she has some medical issue. For our

city we can also store people's identity by their walking postures. Duplicate fingerprints and face reorganization may be possible but no one can change person's walking posture. This feature is unique in all the human beings. Someone who is not from our city so we can identify easily.

Conclusion :

Here it proposed human identification system for a fully convolutional neural network for gait energy image (GEI) reconstruction from an incomplete gait cycle. GEI represents human motion sequence in a single image while preserving temporal information. To overcome incomplete GEI which is composed of only a few frames of a gait cycle the model could reconstruct a GEI. Experimental results on two large gait datasets show that the proposed model can improve recognition rate greatly, particularity when there is only 0.1 part of a gait cycle is available. In future we will extend this model to an end-to-end model for both gait energy image reconstruction and recognition. Having one end-to-end deep model for both GEI reconstruction and recognition is also interesting. In future we are trying to add one more biometric field for human identification. We will work this experiment on dynamic data as a future work.

.

REFERENCES

[1] Anita Thengade and Rucha Dondal(2012), " Recognition of Human Gait Using Genetic Fuzzy Finite State Machine", International Journal of Engineering Research and Technology Vol. 1 (02), ISSN 2278 – 0181.

[2] X. Xing, K. Wang, T. Yan, Z. Lv, Complete canonical correlation analysis with application to multi-view gait recognition, Pattern Recognition 50 (2016) 107–117.

[3] T. Wolf, M. Babaee, G. Rigoll, Multi-view gait recognition using 3d convolutional neural networks, in: Proceedings of the International Conference on Image Processing (ICIP), IEEE, 2016, pp. 4165–4169.

[4] M. Babaee, G. Rigoll, (2017) View-invariant gait representation using joint bayesian regularized non-negative matrix factorization, in: The IEEE International Conference on Computer Vision (ICCV)

[5] K. Bashir, T. Xiang, S. Gong, Gait recognition using gait entropy image.

[6] S. Yu, H. Chen, Q. Wang, L. Shen, Y. Huang, (2017) ,Invariant feature extraction for gait recognition using only one uniform model, Neuro computing 239 81–93.

[7] K. Shiraga, Y. Makihara, D. Muramatsu, T. Echigo, Y. Yagi, Geinet: (2016) View-invariant gait recognition using a convolutional neural network, in: Biometrics (ICB), International Conference on, IEEE, 2016, pp. 1–8.

[8] D. Liu, M. Ye, X. Li, F. Zhang, L. Lin, (2016) Memory-based gait recognition, in: British Machine Vision Conference (BMVC).

[9] M. Alotaibi, A. Mahmood, (2017) Improved gait recognition based on specialized deep convolutional neural network, Computer Vision and Image Understanding 164 103–110.

[10] M. Sajjadi, B. Scholkopf, M. Hirsch, Enhancenet (2017): Single image super-resolution through automated texture synthesis, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, , pp. 4491–4500.

[11] A. Kappeler, S. Yoo, Q. Dai, A. Katsaggelos, (2016) Video super-resolution with convolutional neural networks, IEEE Transactions on Computational Imaging 2 (2) 109–122.

[12] J. Dai, Y. Li, K. He, J. Sun, R-fcn (2016): Object detection via region-based fully convolutional networks, in: Advances in neural information processing systems, , pp. 379–387.

[13] C. Dong, C. C. Loy, K. He, X. Tang, (2016) Image super-resolution using deep convolutional networks, IEEE transactions on pattern analysis and machine intelligence 38 (2) 295–307.